A Cholesky-Based SGM-MLFMM for Stochastic Full-Wave Problems Described by Correlated Random Variables

Zdravko Zubac, Luca Daniel, *Member, IEEE*, Daniel De Zutter, *Fellow, IEEE*, and Dries Vande Ginste, *Senior Member, IEEE*

***Abstract*—In this letter, the multilevel fast multipole method (MLFMM) is combined with the polynomial chaos expansion (PCE)-based stochastic Galerkin method (SGM) to stochastically model scatterers with geometrical variations that need to be described by a set of *correlated* random variables (RVs). It is demonstrated how Cholesky decomposition is the appropriate choice for the RVs transformation, leading to an efficient SGM-MLFMM algorithm. The novel method is applied to the uncertainty quantification of the currents induced on a rough surface, being a classic example of a scatterer described by means of correlated RVs, and the results clearly demonstrate its superiority compared to the nonintrusive PCE methods and to the standard Monte Carlo method.**

***Index Terms*—Cholesky uncertainty quantification (UQ), correlation scattering, method of moments (MoM), multilevel fast multipole method (MLFMM), rough surface, stochastic Galerkin method (SGM).**

I. Introduction

ELECTROMAGNETIC simulation of objects prone to variability has become an important issue. Often, uncertainty quantification (UQ) relies on Monte Carlo (MC) analysis, which requires many calls to a standard deterministic (full-wave) solver, making it not tractable. Recently, the polynomial chaos expansion (PCE) approach was introduced and combined with known computational electromagnetics methods, both in an intrusive and a nonintrusive way [1], [2]. For the scattering analysis of large structures, the multilevel fast multipole method (MLFMM) was combined with the PCE-based stochastic Galerkin method (SGM) [3]. Parallelization of the SGM-MLFMM even led to the efficient UQ of large optical systems [4]. Yet, only variability described by the *independent* RVs could be treated with this method. However, problems affected by variability, e.g., introduced by the manufacturing process, can most often only be described by a set of *correlated* RVs, rather than the independent ones. Then, traditionally, this set of correlated RVs is transformed into a set of independent RVs via the well-known Karhunen–Loève (KL) transformation [5]. Unfortunately, when in the space of the correlated RVs, the so-called correlation length is small, then the total number of independent RVs after KL transform stays as large as the number of correlated RVs, leading to a high-dimensional problem. In [6], where a finite element method was adopted, this was dealt with by dividing the space of variables into subspaces with a correlation length comparable to their size. Nevertheless, when using an integral equation (IE) formulation, where the electromagnetic behavior is described globally, such an approach as described in [6] is not possible. Therefore, in this letter, we introduce another transformation to tackle the correlation, i.e., the Cholesky transformation. This alleviates the curse of dimensionality within the IE-based SGM-MLFMM framework.

This letter is organized as follows. Section II describes the theoretical framework of the stochastic MLFMM with the correlated RVs. An illustrative numerical example of the scattering at a two-dimensional (2-D) rough surface is given in Section III. Section IV concludes the letter.

II. Cholesky-Based SGM-MLFMM

As a generic example for full-wave stochastic problems with correlated RVs, in this letter, we consider 2-D frequency domain scattering from a perfect electrically conducting (PEC) plate of width w, residing in free space. As depicted in Fig. 1, the plate's roughness is stochastically defined by letting the height of M nodes, equidistantly spaced along the x-axis, vary randomly. These heights are described by a set of M correlated Gaussian variables, collected in vector $\boldsymbol{h} = [h_1, h_2, \ldots, h_M]$, and with correlation matrix $\overline{\boldsymbol{\Sigma}}$. The elements of the correlation matrix are given by

$$\Sigma_{ij} = \sigma^2 \exp\left(-\frac{|x_i - x_j|^2}{L_c^2}\right), \quad i, j = 1, \ldots, M \tag{1}$$

where σ is the standard deviation and L_c is the correlation length. Traditionally, in order to apply the PCE, the correlated RVs are converted into independent RVs, collected in vector $\boldsymbol{\xi} = [\xi_1, \xi_2, \ldots, \xi_R]$ via the KL transform as follows:

$$\boldsymbol{h} = \boldsymbol{\mu} + \overline{\boldsymbol{U}}\,\overline{\boldsymbol{\Lambda}}^{1/2}\boldsymbol{\xi} \tag{2}$$

Manuscript received May 27, 2016; revised August 5, 2016 and August 23, 2016; accepted August 23, 2016. Date of publication August 26, 2016; date of current version April 3, 2017. This work was supported by the Interuniversity Attraction Poles Program initiated by the Belgian Science Policy Office.

Z. Zubac, D. De Zutter, and D. Vande Ginste are with the IBCN/Electromagnetics Group, Department of Information Technology, Ghent University/iMinds, Gent 9052, Belgium (e-mail: zdravko.zubac@ugent.be; daniel.dezutter@intec.ugent.be; vdginste@intec.ugent.be).

L. Daniel is with the Computational Prototyping Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139-4307 USA (e-mail: luca@mit.edu).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LAWP.2016.2603232

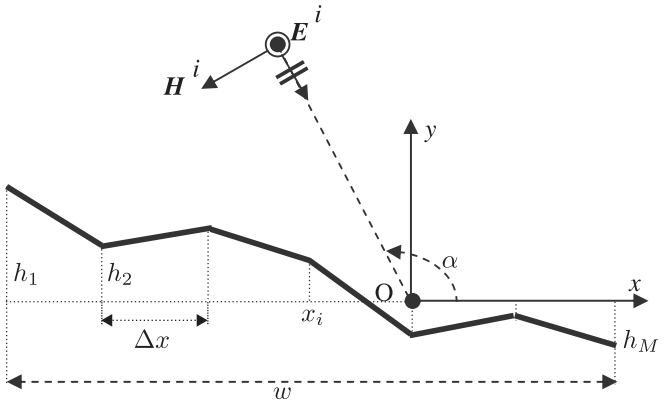

Fig. 1. Rough surface described by a set of correlated random variables (RVs) h_i.

where $\boldsymbol{\mu}$ is the mean value of $\boldsymbol{h}$, and $\overline{\boldsymbol{U}}$ and $\overline{\boldsymbol{\Lambda}}$ are the matrices defined by the eigenvalue decomposition of the correlation matrix $\overline{\boldsymbol{\Sigma}}$, i.e.,

$$\overline{\boldsymbol{\Sigma}} = \overline{\boldsymbol{U}}\,\overline{\boldsymbol{\Lambda}}\,\overline{\boldsymbol{U}}^{T}. \tag{3}$$

Note that the number of independent parameters R may be smaller than the number of correlated parameters M ($R \leqslant M$). The standard electric field IE description of the scattering problem shown in Fig. 1 in conjunction with the method of moments (MoM) yields a linear system that is dependent on $\boldsymbol{\xi}$ [4]

$$\overline{\boldsymbol{Z}}(\boldsymbol{\xi})\boldsymbol{I}(\boldsymbol{\xi}) = \boldsymbol{V}(\boldsymbol{\xi}) \tag{4}$$

where $\overline{\boldsymbol{Z}}(\boldsymbol{\xi})$ is the MoM system matrix, $\boldsymbol{I}(\boldsymbol{\xi})$ is the vector collecting the unknown current densities, and $\boldsymbol{V}(\boldsymbol{\xi})$ is the known right-hand side. All quantities in (4) are expressed in the PCE form, e.g., for $\overline{\boldsymbol{Z}}(\boldsymbol{\xi})$ given as follows:

$$\overline{\boldsymbol{Z}}(\boldsymbol{\xi}) = \sum_{k=0}^{K} \overline{\boldsymbol{Z}}_k\, \phi_k(\boldsymbol{\xi}) \tag{5}$$

where $\{\phi_k(\boldsymbol{\xi})\}_{k=0,\ldots,K}$ represents a set of $K+1$ mutually orthonormal multivariate polynomials according to the Wiener–Askey scheme. In the case of Gaussian variables $\boldsymbol{h}$ (and thus $\boldsymbol{\xi}$), these are products of univariate Hermite polynomials, dependent on a single RV ξ_i. The total number of polynomials grows rapidly with R as

$$K+1 = \frac{(R+P)!}{R!P!} \tag{6}$$

where P is the total order of the polynomials $\phi_k(\boldsymbol{\xi})$, calculated as the sum of the orders of the univariate polynomials they are composed of. Calculation of the PCE coefficients $\overline{\boldsymbol{Z}}_k$ is done via projection, necessitating a multidimensional integration in the R-dimensional space of $\boldsymbol{\xi}$ as follows:

$$\begin{aligned}\overline{\boldsymbol{Z}}_k =\ & < \overline{\boldsymbol{Z}}(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi}) \\ >=\ & \int_{\xi_1} \ldots \int_{\xi_R} \overline{\boldsymbol{Z}}(\boldsymbol{\xi})\, \phi_k(\boldsymbol{\xi})\, W(\boldsymbol{\xi}) d\xi_1 \ldots d\xi_R\end{aligned} \tag{7}$$

where $W(\boldsymbol{\xi})$ represents the multivariate Gaussian probability density function of $\boldsymbol{\xi}$. In particular, when the correlation length L_c is low, the KL transform may lead to a dense, square matrix $\overline{\boldsymbol{U}}\,\overline{\boldsymbol{\Lambda}}^{1/2}$, i.e., $R = M$ and each correlated RV h_i is dependent on all RVs $\boldsymbol{\xi}$. Moreover, each matrix element of $\overline{\boldsymbol{Z}}(\boldsymbol{\xi})$ will also depend on all RVs $\boldsymbol{\xi}$, and the multidimensional integrals of type (7) become cumbersome to compute.

After calculating the coefficients $\boldsymbol{V}_k$ in a similar way, solution of the system (4), for the unknown coefficients $\boldsymbol{I}_k$, is obtained via a Galerkin projection as

$$\boldsymbol{V}_m = \sum_{\substack{k,l=0 \\ \gamma_{klm}\neq 0}}^{K} \overline{\boldsymbol{Z}}_k\, \boldsymbol{I}_l\, \gamma_{klm}, \quad m = 0, \ldots, K \tag{8}$$

where γ_{klm} represents a three-term inner product of Hermite polynomials

$$\gamma_{klm} = \langle \phi_k(\boldsymbol{\xi})\, \phi_l(\boldsymbol{\xi}), \phi_m(\boldsymbol{\xi}) \rangle. \tag{9}$$

Note that (8) constitutes a deterministic linear system with a complexity that scales with the number of nonzero numbers γ_{klm}, which follows an $O(K^{1.5})$ law.

To expedite the solution of the linear system, MLFMM [7] is invoked by dividing the structure into groups of sources. If the distance between a source and an observation group is large enough, then the system (4) can be approximated as

$$\overline{\boldsymbol{D}}(\boldsymbol{\xi})\,\overline{\boldsymbol{T}}\,\overline{\boldsymbol{A}}(\boldsymbol{\xi})\,\boldsymbol{I}(\boldsymbol{\xi}) \approx \overline{\boldsymbol{Z}}(\boldsymbol{\xi}) \tag{10}$$

where $\overline{\boldsymbol{D}}(\boldsymbol{\xi})$, $\overline{\boldsymbol{T}}$, and $\overline{\boldsymbol{A}}(\boldsymbol{\xi})$ represent the well-known disaggregation, translation, and aggregation matrix, respectively. However, in contrast to the problems described in [3], whereas the aggregation and disaggregation matrices were dependent only on a group of sources, and thus only on few h_i, here they are still dependent on all independent RVs $\boldsymbol{\xi}$. Besides the aforementioned curse of dimensionality in calculating PCE projections (7), this also entails an unacceptably long solution time of (8). Indeed, since the aggregation and the disaggregation matrices are dependent on all independent RVs, their PCE coefficients are all nonzero, and the complexity does not scale linearly with the number of polynomials K as in [3], but with the total number of γ_{klm}.

To tackle this issue, instead of using the traditional KL transform, we propose to adopt a Cholesky transformation. Then, the correlated RVs are expressed via another vector of independent RVs $\boldsymbol{\eta}$

$$\boldsymbol{h} = \boldsymbol{\mu} + \overline{\boldsymbol{L}}\,\boldsymbol{\eta} \tag{11}$$

where $\overline{\boldsymbol{L}}$ is a lower triangular matrix related to the correlation matrix as follows [8]:

$$\overline{\boldsymbol{\Sigma}} = \overline{\boldsymbol{L}}\,\overline{\boldsymbol{L}}^{T}. \tag{12}$$

To show the benefits of this Cholesky decomposition, for a canonical structure as shown in Fig. 1, with $M = 200$, $L_c = \lambda/5, \sigma = \lambda/20$, and $w = 20\lambda$ (with λ the free-space wavelength), we present the structure of this particular correlation matrix in Fig. 2 and its corresponding KL and Cholesky matrices in Fig. 3. Whereas the KL matrix $\overline{\boldsymbol{U}}\,\overline{\boldsymbol{\Lambda}}^{1/2}$ is a densely filled matrix, the off-diagonal elements of the Cholesky matrix $\overline{\boldsymbol{L}}$ rapidly vanish as can be seen in Fig. 3. As of yet, a formal proof of this behavior is still missing.

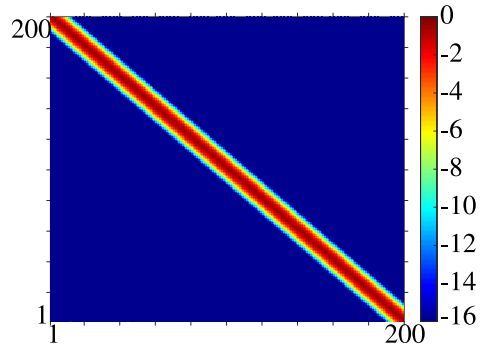

Fig. 2. Magnitude (on a logarithmic scale) of the elements of the correlation matrix $\overline{\boldsymbol{\Sigma}}$ for a canonical problem.

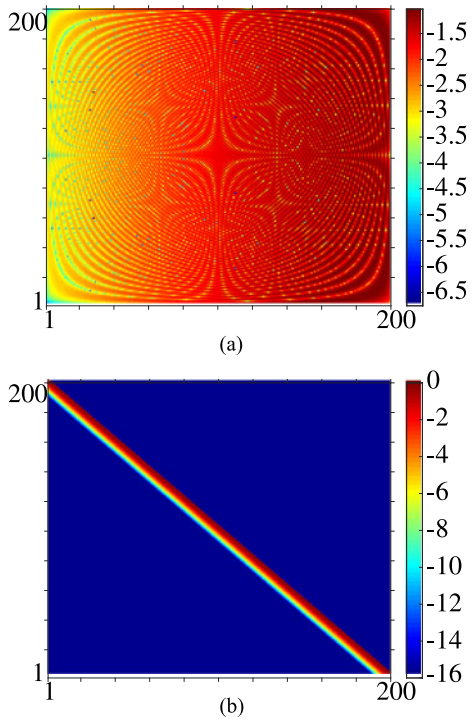

Fig. 3. Magnitude (on a logarithmic scale) of the elements the matrices pertaining to the decomposition of $\overline{\boldsymbol{\Sigma}}$ as shown in Fig. 2. (a) KL matrix. (b) Cholesky matrix.

Consequently, when dimensionality reduction with KL transform is not possible, the benefits of the advocated Cholesky approach are as follows.

1) The M correlated RVs $\boldsymbol{h}$ depend only on a few independent RVs $\boldsymbol{\eta}$. Thus, the M-dimensional integrals of type (7) depending on these correlated RVs are reduced in dimension, and their computation is expedited.
2) Many PCE coefficients are zero, as their corresponding stochastic quantities, in particular the elements of $\overline{\boldsymbol{Z}}(\boldsymbol{\eta})$, only depend on a few independent RVs. This substantially improves the computational and memory complexity.

III. Numerical Example

We consider scattering from a rough PEC strip of width $w = 100\,\lambda$, whose roughness is described by 81 RVs that determine the y-coordinates of the equally distributed points on the structure, as presented in Fig. 1. The correlation length is $L_c = \lambda$, and the standard deviation is $\sigma = \lambda/20$. The incident field is a TM-polarized plane wave impinging under an angle of $\alpha = 3\pi/4$. The structure is discretized with $N = 2000$ segments, and the unknown current density is defined by adopting piecewise constant basis functions. All computations are carried out on a Dell PC with a quad-core Intel Core i7-2600 processor operating at 3.40 GHz and with 8 GB RAM.

TABLE I
Setup and Solution Time

Method	Setup	Solution
SGM	99 s	3948 s
SCM	26 570 s	7971 s
MC	19 200s	5887 s

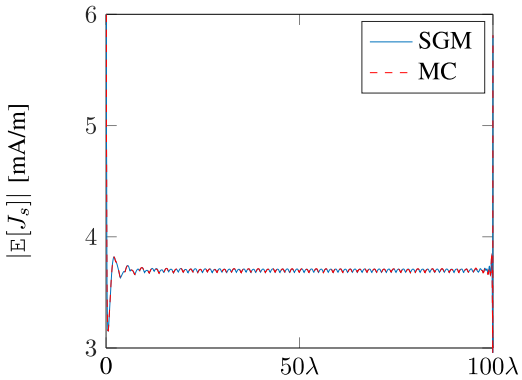

Fig. 4. Average current density $\mathrm{E}[J_s]$ on the rough strip, with $\mathrm{E}[\cdot]$ the expectation operator.

To validate the accuracy and demonstrate the efficiency of our novel method, a standard MC analysis with 10 000 samples is used as a reference solution. This MC analysis takes about 7 h. Moreover, the stochastic scattering problem is also solved by means of the *non*-intrusive PCE-based stochastic collocation method (SCM) leveraging sparse Smolyak integration [9]. The results for the average current density with the SGM-MLFMM scheme and polynomial order 2 lead to an accuracy of 0.15% compared to MC. The accuracy of SGM-MLFMM cannot be readily predicted beforehand, but it can be increased by increasing the polynomial order. Also, SCM uses $P = 2$ and 13 285 Smolyak integration points. The timing analysis is as follows: The novel SGM-MLFMM scheme takes about 1 h, and SCM takes around 10 h. The gain is achieved in both the setup and the solution phase, as is visible from Table I. If the KL transform would be used in combination with SGM, then the setup time for calculation of the PCE coefficients of $\overline{\boldsymbol{D}}$ and $\overline{\boldsymbol{A}}$ coefficients would be determined by (7), and would be of the same order of magnitude as the setup time of SCM.

The average current density on the strip is given in Fig. 4, and its standard deviation is presented in Fig. 5. Good agreement between SGM and MC is visible. These results are presented for polynomial order $P = 2$ and the corresponding total number of stochastic unknowns $N_{\text{stoc}} = (K + 1)N = 6806\,000$.

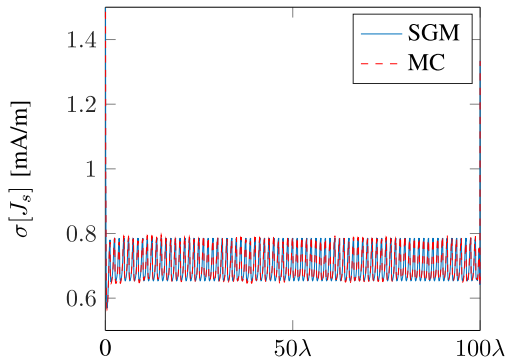

Fig. 5. Standard deviation of the current density $\sigma[J_s]$ on the rough strip, with $\sigma[J_s] = \sqrt{E[(J_s - E[J_s])(J_s - E[J_s])^H]}$.

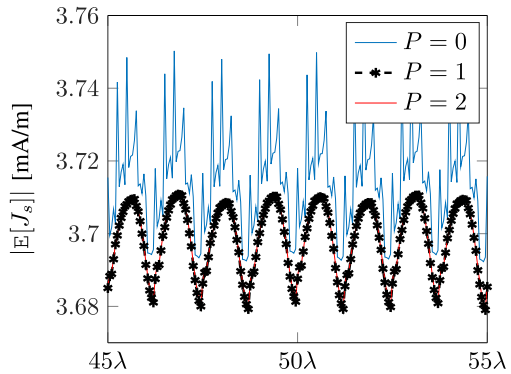

Fig. 6. Average current density in the middle of the rough strip for several polynomial orders.

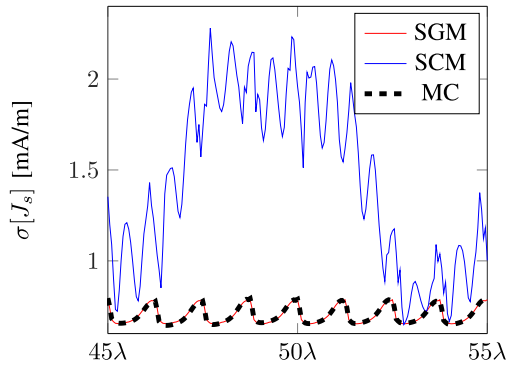

Fig. 7. Standard deviation of the current density J_s in the middle of the rough strip.

To reduce possible truncation errors, as described in [3], the polynomial order should be chosen large enough such that the PCE of $\overline{\boldsymbol{Z}}$ can be found accurately through multiplication and Galerkin projection of the PCE coefficients of $\overline{\boldsymbol{D}}$ and $\overline{\boldsymbol{A}}$. To demonstrate the influence of the truncation error on the average current density, in Fig. 6 we present $\mathrm{E}[J_s]$ in the middle of the strip for several polynomial orders. From this figure, the convergence of the advocated SGM-MLFMM scheme is clearly visible, which also again validates our method. Moreover, at this point, it is important to point out that, in particular, when dealing with full-wave problems, variations of the output parameters, such as current density, can be substantial and the Smolyak integration rule used in SCM may fail to produce good results. This is visible from Fig. 7, where the standard deviation of the current density J_s in the middle of the rough strip is shown. This behavior is well known for the integration of functions that are not smooth enough [10]. The proposed SGM-MLFMM does not suffer from this issue, however, since the integration was done in a lower dimensional space thanks to the advocated Cholesky transformation. To achieve the same level of accuracy for the standard deviation, with the SGM, the number of Smolyak integration points should be increased to 722 089, which becomes prohibitively expensive. This clearly demonstrates the huge advantage of the novel SGM-MLFMM scheme over SCM.

IV. Conclusion

In this letter, the UQ of full-wave stochastic problems, described by *correlated* RVs, was investigated. Classically, the KL transformation is applied to decorrelate the RVs. However, for the envisaged applications, the SGM-MLFMM scheme, presented in the literature before by the authors, cannot be straightforwardly extended by incorporating a KL transformation, and this is because of two reasons: 1) the computation of the PCE coefficients entails integration in a highly dimensional space; and 2) all these PCE are nonzero, as the stochastic quantities are dependent on all independent RVs after KL transformation. We proposed to tackle these issues by invoking Cholesky decomposition of the correlation matrix instead, leading to a very accurate and efficient SGM-MLFMM algorithm. The novel method was validated and compared to an MC analysis and an SCM for the case of scattering at rough PEC plate.

References

[1] T. El-Moselhy and L. Daniel, "Variation-aware stochastic extraction with large parameter dimensionality: Review and comparison of state of the art intrusive and non-intrusive techniques," in *Proc. 12th Int. Symp. Quality Electron. Des.*, Santa Clara, CA, USA, Mar. 2011, pp. 508–517.

[2] C. Chauvière, J. S. Hesthaven, and L. C. Wilcox, "Efficient computation of RCS from scatterers of uncertain shapes," *IEEE Trans. Antennas Propag.*, vol. 55, no. 5, pp. 1437–1448, May 2007.

[3] Z. Zubac, D. De Zutter, and D. Vande Ginste, "Scattering from two-dimensional objects of varying shape combining the multilevel fast multipole method (MLFMM) with the stochastic Galerkin method (SGM)," *IEEE Antennas Wireless Propag. Lett.*, vol. 13, pp. 1275–1278, 2014.

[4] Z. Zubac, J. Fostier, D. De Zutter, and D. Vande Ginste, "Efficient uncertainty quantification of large two-dimensional optical systems with a parallelized stochastic Galerkin method," *Opt. Express*, vol. 23, no. 24, pp. 30 833–30 850, Nov. 2015.

[5] M. Loeve, *Probability Theory II*. New York, NY, USA: Springer-Verlag, 1994.

[6] Y. Chen, J. Jakeman, C. Gittelson, and D. Xiu, "Local polynomial chaos expansion for linear differential equations with high dimensional random inputs," *SIAM J. Sci. Comput.*, vol. 37, no. 1, pp. A79–A102, 2015.

[7] W. C. Chew, J. M. Jin, E. Michielssen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*. Norwood, MA, USA: Artech House, 2001.

[8] N. J. Higham, "Cholesky factorization," *WIREs Comput. Stat.*, vol. 1, no. 2, pp. 251–254, 2009.

[9] Z. Zubac, D. De Zutter, and D. Vande Ginste, "Scattering from two-dimensional objects of varying shape combining the method of moments with the stochastic Galerkin method," *IEEE Trans. Antennas Propag.*, vol. 62, no. 9, pp. 4852–4856, Sep. 2014.

[10] P. Tsuji, D. Xiu, and L. Ying, "Fast method for high-frequency acoustic scattering from random scatterers," *Int. J. Uncertainty Quantification*, vol. 1, no. 2, pp. 99–117, 2011.